

On Information Metrics for Spatial Coding

Bryan C. Souza,^{a,*†} Rodrigo Pavão,^{a,†} Hindiael Belchior^b and Adriano B. L. Tort^a

^a Brain Institute, Federal University of Rio Grande do Norte, RN, Brazil

^b Faculty of Health Sciences of Trairi, Federal University of Rio Grande do Norte, RN, Brazil

Abstract—The hippocampal formation is involved in navigation, and its neuronal activity exhibits a variety of spatial correlates (e.g., place cells, grid cells). The quantification of the information encoded by spikes has been standard procedure to identify which cells have spatial correlates. For place cells, most of the established metrics derive from Shannon's mutual information (Shannon, 1948), and convey information rate in bits/s or bits/spike (Skaggs et al., 1993, 1996). Despite their widespread use, the performance of these metrics in relation to the original mutual information metric has never been investigated. In this work, using simulated and real data, we find that the current information metrics correlate less with the accuracy of spatial decoding than the original mutual information metric. We also find that the top informative cells may differ among metrics, and show a surrogate-based normalization that yields comparable spatial information estimates. Since different information metrics may identify different neuronal populations, we discuss current and alternative definitions of spatially informative cells, which affect the metric choice. © 2018 IBRO. Published by Elsevier Ltd. All rights reserved.

Key words: place cell, place field, spatial coding, information, spike train analysis, hippocampus.

INTRODUCTION

The hippocampus is known to be involved in memory formation (Scoville and Milner, 1957; Eichenbaum, 2000; Zola-Morgan and Squire, 1990) and spatial navigation (O'Keefe and Dostrovsky, 1971; Morris et al., 1982; O'Keefe and Recce, 1993). In the early 70s, O'Keefe and Dostrovsky discovered that some hippocampal cells have their firing rate modulated by the animal position, discharging more at a spatial region known as the place field of the cell (O'Keefe and Dostrovsky, 1971). Since the discovery of place cells, other types of spatial correlates emerged in areas related to the hippocampal circuitry, such as the head-direction cells in the postsubiculum (Taube et al., 1990), and the grid cells and speed cells in the entorhinal cortex (Fyhn et al., 2004; Hafting et al., 2005; Moser et al., 2008; Kropff et al., 2015). Properly identifying these cells requires estimates of the information contained in spikes about navigational features (i.e., position, speed, head angle). The main metrics used to estimate this type of information were proposed by Skaggs et al. (1993, 1996) and are derivations from Shannon's mutual information (MI).

Information entropy, as originally proposed by Shannon, measures the amount of uncertainty in the outcome of a variable based on its probability of occurrence (Shannon, 1948). In other words, the more unpredictable the outcome is, the more entropy it has. On the other hand, the MI is a measure of the shared entropy between two variables; it indicates how much knowing a variable X reduces the uncertainty of a variable Y. While the MI is usually measured in bits, the two derived metrics proposed by Skaggs et al. express information in bits per spike (I_{spike}) or bits per second (I_{sec}). These metrics are not straightforward divisions of the MI by the number of spikes or elapsed time, but are rather defined as estimates of the average information rate conveyed by the cell (Skaggs et al., 1993); this is achieved by keeping the lower order terms of the MI power series expansion with respect to time (see Methods). A fundamental difference from the MI is that I_{sec} and I_{spike} only take into account the average firing rate over the spatial variable (e.g., location, speed), ignoring firing rate differences across multiple occurrences of the same variable (e.g., across trials). Although these metrics provide a meaningful interpretation of the relation between firing rate and navigational features, the possible implications introduced by these modifications remain to be investigated. Of note, the MI has been previously used to measure spatial information (Ego-Stengel and Wilson, 2007), though a direct comparison with the I_{sec} and I_{spike} metrics has never been performed.

*Corresponding author.

E-mail address: bryanacsouza@neuro.ufrn.br (B. C. Souza).

† Current address: Department of Neuroscience, Albert Einstein College of Medicine, NY, USA.

‡ Equal contribution.

Abbreviation: MI, mutual information.

In this work, we use simulated data to address how well the metrics I_{spike} , I_{sec} and MI reflect the capacity of decoding the animal position based on the spikes of an individual neuron, which directly relates to the amount of spatial information conveyed by the cell (Quiroga and Panzeri, 2009). We find that while MI values correlate well with decoding performance under a variety of scenarios, this is not always the case for I_{spike} and I_{sec} . Similar results hold when analyzing real spikes from place cells of rats recorded on a linear track. Moreover, we also find that the different metrics may give rise to different subpopulations of spatially modulated neurons. Finally, we show that a surrogate-based normalization can equalize the three metrics. We end by discussing the conceptual definition of spatially informative cells, which may vary according to the employed metric.

METHODS

Simulating spatially modulated cells

We simulated the firing rate of spatially modulated neurons across 30 trials on a linear track divided into 25 bins of space (Fig. 1A). For simplicity, the animal speed and occupancy were considered constant over space. We modeled eight types of cells; for each cell type, we simulated 10 levels of spatial modulation (Fig. 1B, Neuron ID a to j). For the first 5 cell types (Neuronal types I–V), the firing rate of each trial was modeled as a Gaussian centered (on average) at bin 13 with (average) standard deviation of 5 bins. To introduce inter-trial variability, 0.5 and 0.1 white noise was added

to the center and standard deviation of the Gaussian, respectively. The cell types mimicked the behavior of (I) a pyramidal-like place cell (low basal firing rate); (II) an interneuron-like place cell (high basal firing rate); (III) interneuron-like place cell negatively modulated by space; (IV) a pyramidal-like place cell exhibiting spatial modulation only in a subset of trials; and (V) a cell as in IV, but with constant mean firing rate across trials. For cell types I–III, we varied spatial modulation strength (deviation from baseline), while for cell types IV and V we varied the number of modulated trials. Cell type VI was similar to cell type I but could have multiple, equally spaced peaks as a grid cell. We also simulated cells behaving as ramp (VII) or constant functions (VIII) along space, with different slopes and firing rate levels, respectively.

Estimating spatial information in the firing rate

To estimate the spatial information contained in the firing rate of each cell, we used I_{spike} and I_{sec} – the main metrics for selecting place cells (Skaggs et al., 1993, 1996) – and the MI (Shannon, 1948). We computed the I_{sec} metric from the average firing rate (over trials) in the 25 space bins using the following definition:

$$I_{\text{sec}} = \sum_{i=1}^{25} p_i \lambda_i \log_2 \frac{\lambda_i}{\lambda}$$

where λ_i is the mean firing rate in the i -th space bin and p_i the occupancy ratio of the bin, while λ is the overall mean firing rate of the cell. I_{sec} measures information rate in bits per second (see next section for mathematical derivation).

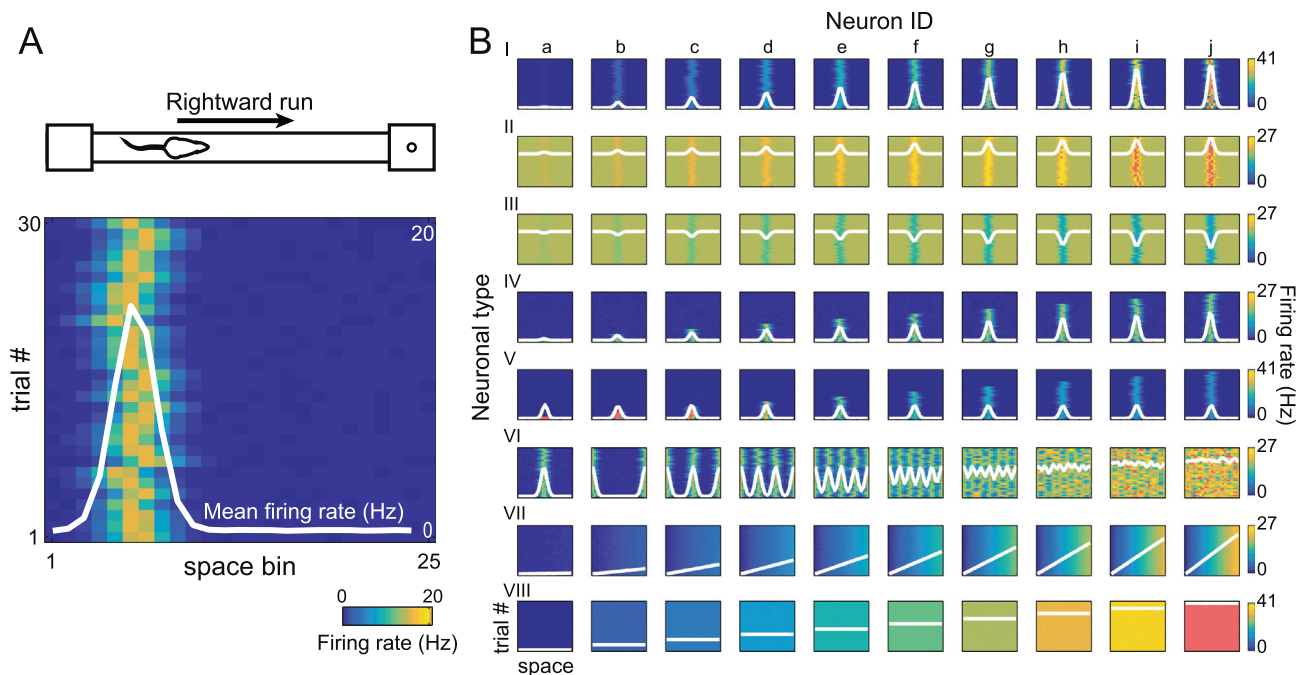


Fig. 1. Simulating the activity of spatially modulated neurons. (A) Neuronal firing rate was simulated in 25 bins of space over 30 trials, meant to represent spiking activity of one cell during rightward runs on a linear track. In the rate-position map, pseudo-colors represent firing rate on each trial, while the white line is the mean firing rate over trials. (B) Rows display 8 simulated neuronal types (I–VIII) of different spatial modulation profiles. Cell types I–V exhibit place-cell-like behavior, while types VI, VII and VIII represent grid cells, ramp and constant functions, respectively (see Methods for details). For each type, columns show the rate-position maps (pseudo-color scale) as well as the mean firing rate (overlying white trace) for 10 model cells (a–j) differing in the level of spatial modulation.

The I_{spike} metric is a normalization of I_{sec} , defined as:

$$I_{\text{spike}} = \frac{I_{\text{sec}}}{\lambda}$$

This normalization yields values in bits per spike.

The MI was estimated using all firing rate values (within trials), which were binned into four non-overlapping quantiles:

$$MI = \sum_{i=1}^{25} \sum_{j=1}^4 p_{ij} \log_2 \left(\frac{p_{ij}}{p_i \cdot p_j} \right)$$

where p_i and p_j are the probabilities of position bin i and firing rate bin j , respectively; p_{ij} is the joint probability between position bin i and firing rate bin j .

Derivation of I_{sec} from MI

Using Bayes' rule, the MI can be rewritten as:

$$MI = \sum_i \sum_j p_{ji} p_i \log_2 \left(\frac{p_{ji}}{p_j} \right)$$

where p_{ji} is the conditional probability of firing rate bin j given the position i . For a sufficient small amount of time Δt , we can assume that a cell can either emit just one spike or none, and thus there are only two possible firing rate bins, denoted as $j = 0$ (no spike) and $j = 1$ (spike). The probability of spike occurrence during Δt at position i is given by $p_{j=1|i} = \lambda_i \Delta t$, while $p_{j=0|i} = 1 - \lambda_i \Delta t$ denotes the probability of no spike. Similarly, $p_{j=1} = \lambda \Delta t$ and $p_{j=0} = 1 - \lambda \Delta t$ denote spiking probabilities irrespective of position. Using these probabilities in the MI formula above gives:

$$\begin{aligned} MI &= \sum_i \lambda_i \Delta t p_i \log_2 \left(\frac{\lambda_i \Delta t}{\lambda} \right) + (1 - \lambda_i \Delta t) p_i \log_2 \left(\frac{1 - \lambda_i \Delta t}{1 - \lambda \Delta t} \right) \\ &= \sum_i \lambda_i \Delta t p_i \log_2 \left(\frac{\lambda_i}{\lambda} \right) + p_i \log_2 \left(\frac{1 - \lambda_i \Delta t}{1 - \lambda \Delta t} \right) - \lambda_i \Delta t p_i \log_2 \left(\frac{1 - \lambda_i \Delta t}{1 - \lambda \Delta t} \right) \\ &= \sum_i \lambda_i \Delta t p_i \log_2 \left(\frac{\lambda_i}{\lambda} \right) + p_i \log_2 (1 - \lambda_i \Delta t) - p_i \log_2 (1 - \lambda \Delta t) \\ &\quad - \lambda_i \Delta t p_i (\log_2 (1 - \lambda_i \Delta t) - \log_2 (1 - \lambda \Delta t)) \end{aligned}$$

Using the first term in the power series expansion of logarithms, we have:

$$\log_2 (1 - x \Delta t) \approx \frac{-x \Delta t}{\ln(2)}$$

Applying this approximation in the previous MI equation yields:

$$MI \approx \sum_i \lambda_i \Delta t p_i \log_2 \left(\frac{\lambda_i}{\lambda} \right) - \frac{p_i \lambda_i \Delta t}{\ln(2)} + \frac{p_i \lambda \Delta t}{\ln(2)} - \frac{\lambda_i \Delta t p_i (-\lambda_i \Delta t + \lambda \Delta t)}{\ln(2)}$$

Excluding the second order terms (Δt^2), we have:

$$MI \approx \sum_i \lambda_i \Delta t p_i \log_2 \left(\frac{\lambda_i}{\lambda} \right) - \frac{\Delta t}{\ln(2)} \sum_i p_i \lambda_i + \frac{\lambda \Delta t}{\ln(2)} \sum_i p_i$$

Using $\lambda = \sum_i p_i \lambda_i$ and $\sum_i p_i = 1$, the last two terms cancel out, yielding:

$$MI \approx \Delta t \sum_i \lambda_i p_i \log_2 \left(\frac{\lambda_i}{\lambda} \right)$$

Finally, since the power series expansion of MI around t is given by:

$$MI(t + \Delta t) = \sum_{k=0}^{\infty} \Delta t^k \frac{MI^{(k)}}{k!}$$

where $MI^{(k)}$ is the k -th time derivative, we have that the first time derivative is approximated by:

$$MI^{(1)} \approx \sum_i \lambda_i p_i \log_2 \left(\frac{\lambda_i}{\lambda} \right)$$

which is the definition of I_{sec} .

Estimating spatial information from decoding performance

Decoding algorithms predict the most likely stimulus that generated a given response based on the previous observations of stimulus–response pairs (Quiroga and Panzeri, 2009). We used a Gaussian naïve Bayes classifier to predict the position of the animal based on the firing rate of a cell (see John and Langley, 1995 for detailed description). This approach has been previously used in neuroscience research to infer behavioral and sensory correlates of neuronal activity (Brown et al., 1998; Zhang et al., 1998; Salinas and Romo, 1998; Barbieri et al., 2005). It defines a conditional probability model with the prior probability of position and the likelihood probability of firing rate given the position, which is assumed to be normally distributed. These probabilities are estimated based on the available samples and used to compute the posterior probability of position given the firing rate. New firing rate samples can then be assigned to the most probable position as defined by the *maximum a posterior probability*.

We performed the classification using a leave-one-out approach (Fig. 2). Briefly, all firing rate values (across positions and trials; training bins) but one (test bin) are used to estimate the posterior probabilities (Fig. 2A). The model then predicts the position of the left out firing rate value. This procedure is repeated multiple times so that each firing rate value is used once as a test bin. A confusion matrix is next constructed from actual and decoded positions, and used to extract the percentage of correct decoding, which is the proportion of entries in the $y = x$ diagonal (Fig. 2B). In this work, the percentage of correct decoding was assumed to correlate with the true spatial information content of the cell.

Normalizing spatial information metrics using surrogates

To normalize the estimates of spatial information, we first shuffled the labeling of position bins on each trial (Fig. 4A). This approach avoids any firing preference across trials and in the mean rate. We then computed the I_{spike} , I_{sec} and MI metrics using the shuffled bins. This procedure was repeated 100 times and used to build a surrogate distribution for each metric. The

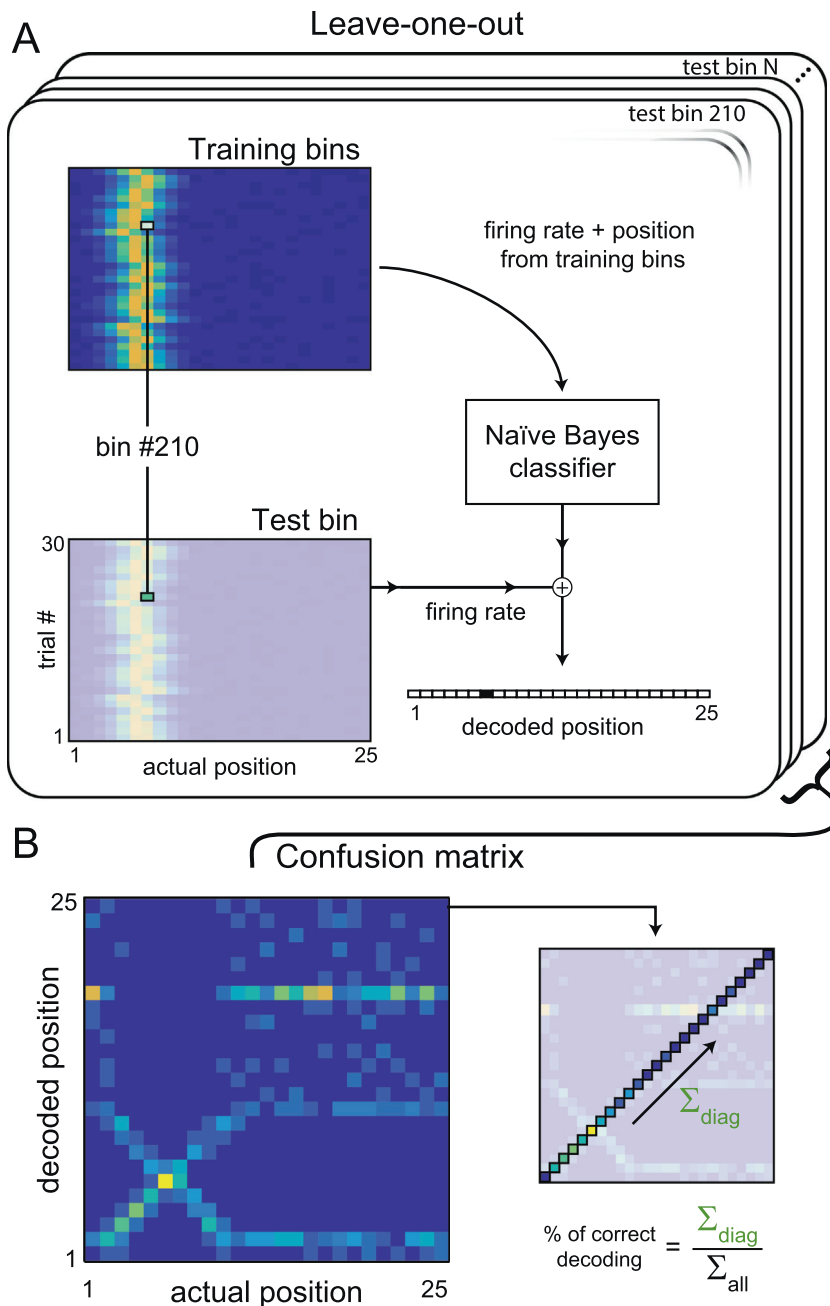


Fig. 2. Estimating spatial information using a Bayesian classifier. (A) All the (training) bins of the rate-position map except one were used to train a naïve Bayes classifier. The classifier was then used to decode the position of the remaining (test) bin using only its firing rate. Each bin was used once as a test bin (leave-one-out approach). (B) All the decoded positions were used to compute the confusion matrix, which relates decoded and actual positions. The number of correct classifications (sum of the matrix diagonal) was divided by the total number of classifications to calculate the decoding performance.

information values were then expressed as z-scores of the surrogate distribution. The z-scored metrics are referred to as normalized I_{spike} (Norm. I_{spike}), I_{sec} (Norm. I_{sec}) and MI (Norm. MI).

Measuring spatial information in real neurons

To investigate the relation between the spatial information estimates and decoding performance for real cells, we

used a dataset with recordings from the CA1 region of the dorsal hippocampus of three rats running back and forth on a linear track (Mizuseki et al., 2013; data freely available at <https://crcns.org/>). We calculated the spatial information estimates and the decoding performance for all neurons in 75 recording sessions. In Fig. 8, we used the classification of putative interneurons and pyramidal cells available in the dataset (Mizuseki et al., 2009).

Comparing the subpopulations of spatially informative neurons

We investigated the overlap between the subpopulations of most informative neurons according to the different metrics. First, for each metric we ranked the cells from most to least informative. We then computed the percentage of common cells between each pairwise combination of metrics for the N% most informative cells, with N% varying from 20% to 100%.

RESULTS

We first computed the three spatial metrics for each of the 8 simulated cell types shown in Fig. 1. These cell types vary in how their firing rate is modulated by space (see Methods). For instance, while cell type I fires at its place field location on every trial, cell type IV emits spatially modulated spikes only in a subset of trials. For each cell type, we varied the amount of spatial modulation in 10 levels. This was achieved by changing the deviation of the firing rate from the basal level (cell types I–III), the percentage of trials with modulated activity (cell types IV and V), the number of place fields (cell type VI), the spatial slope (cell type VII) or the basal firing rate level (cell type VIII). The spatial metrics were then compared with the percentage of correct decoding of a Bayesian classifier (Fig. 2).

Fig. 3 shows examples in which the same decoding performance could have either high or low values of I_{sec} and I_{spike} . For instance, different simulations of cell

type I (classical place cell) at varying levels of spatial modulation – that is, the amount of firing rate deviation from baseline when passing through the place field; see Fig. 3A – revealed that I_{sec} tends to provide increasing levels of information at similar decoding performance, while the MI yielded steadier values (Fig. 3B; compare also the connected red circles in Fig. 3C). On the other hand, increasing the number of trials in which spatial

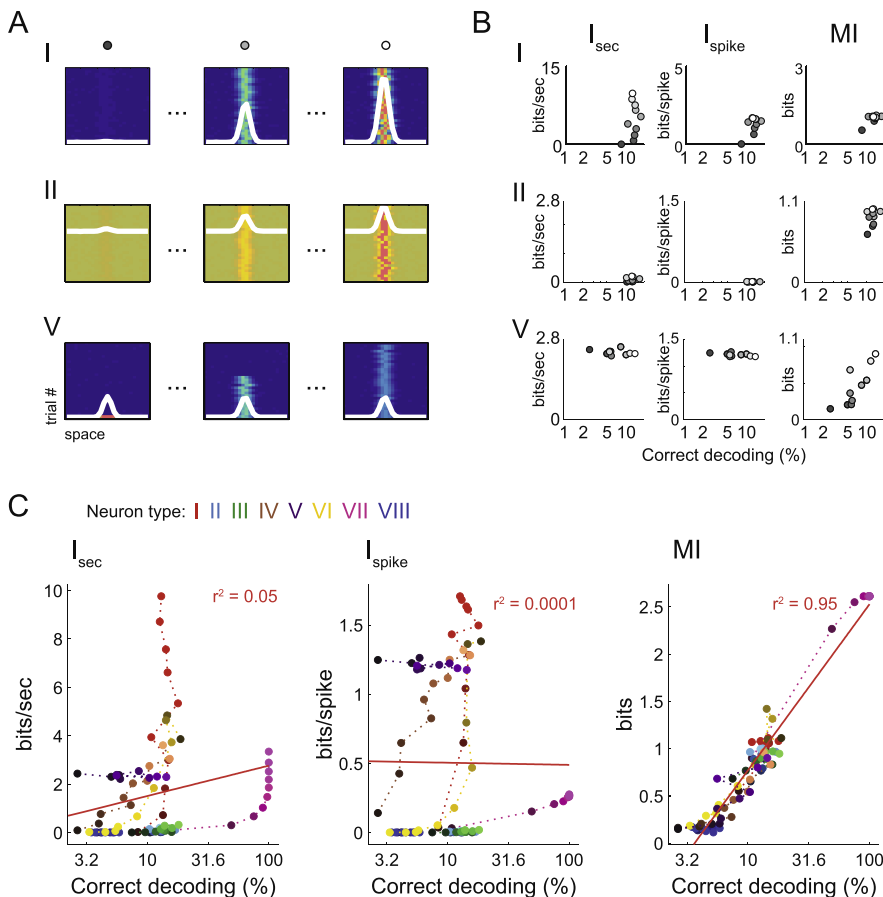


Fig. 3. Assessing the performance of spatial information metrics. (A) Examples of cell types I, II and V during three modulation conditions. Circles on top indicate the strength of spatial modulation, from black (low) to white (high). (B) I_{sec} , I_{spike} and MI values of the cell types in A plotted against the percentage of correct decoding of the animal position (see Methods). Note that the same level of decoding across different cell types can elicit distinct I_{sec} and I_{spike} values (i.e., compare rows), and that similar values of I_{sec} or I_{spike} within a same cell type may be associated to different decoding levels (i.e., see cell type V). (C) Percentage of correct decoding in log scale for all simulated cells vs. I_{sec} , I_{spike} and MI along with the linear fit (red line). Colors denote neuronal types. Dotted line and color gradient go from the lowest (dark) to highest (light) modulation levels (Neuron ID in Fig. 1B). The MI best correlates with decoding performance.

selective firing occurred for cell type V (“trial consistency”) was associated with higher decoding performance and MI values while information values estimated from both I_{sec} and I_{spike} remained roughly constant (Fig. 3B and connected purple circles in Fig. 3C). There were also cases in which, in contrast to the MI, the I_{sec} and I_{spike} estimates were low and did not reflect the capacity of decoding spatial information from the cell (e.g., see cell type II in Fig. 3A, B).

We also note that even though the I_{sec} and I_{spike} tended to increase with decoding performance for some cells (e.g., see cell types IV, VI and VII in Fig. 3C), the estimated information values for a fixed percentage of decoding performance were not consistent across cells. For example, notice in Fig. 3B that a 10% decoding performance was associated with ~2 bits/spike for cell type I, ~0.01 bits/spike for cell type II and ~1.3 bits/spike for cell type V. On the other hand, the MI showed much higher consistency across the different cells types. Accordingly, pooled data from all simulated cells

revealed a clear correlation between the percentage of correct decoding in log scale and MI, but not between correct decoding and either I_{sec} or I_{spike} (Fig. 3C).

We next corrected the spatial information metrics for the chance information level of each cell, which varies according to firing characteristics. To that end, we first shuffled the positions of firing rate bins within trials (Fig. 4A left), and then computed the information metrics using as input data the shuffled rate-position map. This procedure was repeated multiple times to build a distribution of surrogate information values for each metric (Fig. 4A right). Finally, the actual information values were normalized by the chance distribution through z-scoring in relation to the corresponding surrogate mean and standard deviation. Using the same cell types and analyses as in Fig. 3, we found that while the correlation between decoding performance and MI did not considerably change, the correlation of I_{sec} and I_{spike} with decoding substantially improved after the normalization and became similar across metrics (Fig. 4B). Accordingly, both I_{sec} and I_{spike} displayed higher consistency across cell types following the normalization (compare Fig. 3C and Fig. 4B).

We next investigated spikes from real cells recorded from the rat hippocampus during traversals on a linear track. Fig. 5A shows example rate-position maps of five CA1 units with different spatial modulation

profiles that were recorded in the same session. Similar to simulated data, the original and normalized MI, I_{sec} and I_{spike} were computed for each cell, and plotted against decoding performance in log scale. Fig. 5B, C shows scatter plots of information and decoding performance and their correlation computed using all neurons in this example session. Notice in Fig. 5B that the MI values correlated well with decoding performance, while I_{spike} showed no clear correlation (MI: $r^2 = 0.67$; I_{spike} : $r^2 = 0.007$). Unexpectedly, the I_{sec} values were better correlated to decoding than in our simulations (I_{sec} : $r^2 = 0.51$). The correlation with decoding substantially improved for I_{sec} and I_{spike} after the normalization, and reached similar values across all three metrics (Fig. 5C). Fig. 6A shows the r^2 values between each metric and decoding performance for 75 linear track sessions. Consistent with the example in Fig. 5, we found that MI held the best correlations with decoding performance, followed by I_{sec} , while I_{spike} was

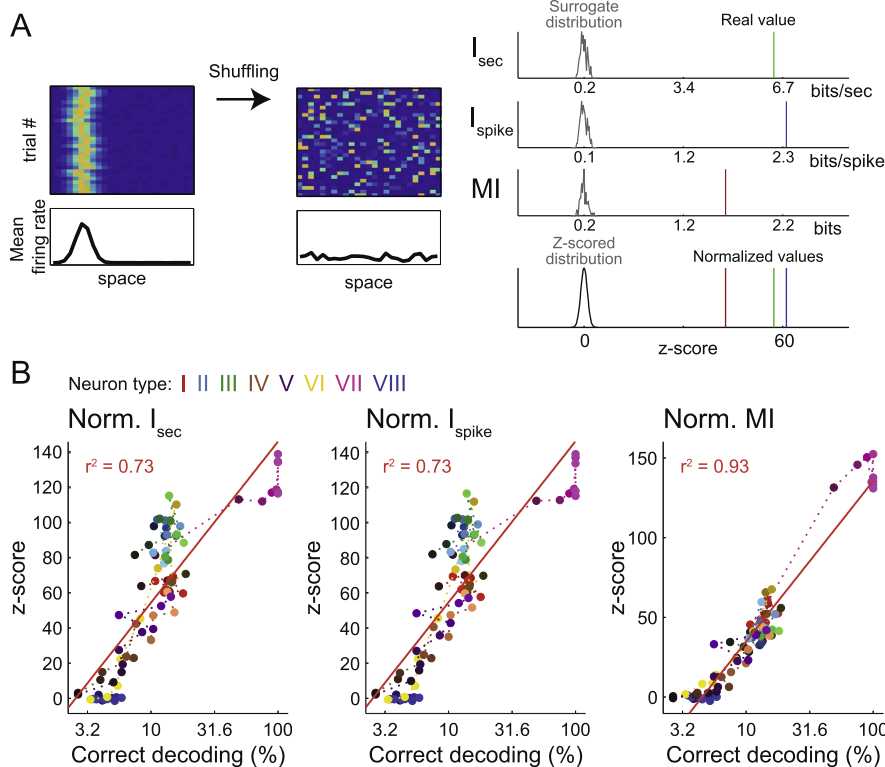


Fig. 4. Correcting spatial metrics using surrogates. (A) (Left) Example of shuffling procedure. To estimate chance information levels, the firing rate bins within trials were shuffled prior to computing the metrics. (Right) The actual value of each metric was compared to the surrogate distribution ($n = 100$ shufflings) and z-score normalized. (B) Percentage of correct decoding (in log scale) vs. the normalized values of I_{sec} , I_{spike} and MI for all simulated cells along with the linear fit (red line). Colors denote neuronal types (as in Fig. 3). Normalizing I_{sec} and I_{spike} significantly improves their correlation with decoding performance.

poorly correlated. After normalization, the three metrics correlated similarly well with decoding. Fig. 6B shows the changes in r^2 between the original and normalized versions of each metric. While the r^2 of I_{sec} and I_{spike} significantly increased, the r^2 changes for MI were not considerably different.

We proceeded to investigate how similar are the subpopulations of spatially selective cells when defined by each metric at varying degrees of information threshold. To that end, for each metric we ranked neurons according to their spatial information (Fig. 7A), and then computed the pairwise intersection (i.e., MI & I_{sec} , MI & I_{spike} , I_{sec} & I_{spike}) between the top N% informative cells. Fig. 7B shows the percentage of common cells as a function of N%. For up to the top 40% spatially informative cells, we found that there is very low intersection for I_{spike} and MI and for I_{spike} and I_{sec} (median below 20% of common cells). As expected, the intersection between metrics gradually increases with decreasing the information threshold (i.e., increasing N%), eventually reaching a complete intersection when all cells are analyzed. On the other hand, MI and I_{sec} had a high number of common cells for all information thresholds, with a median intersection of ~90% of cells. Therefore, the subpopulation of spatially informative neurons according to the MI is slightly different from I_{sec} and drastically different from

I_{spike} (Fig. 7B). Surprisingly, I_{sec} and I_{spike} showed very low overlap among their top informative cells despite their related definitions. We also computed the rank correlation between the metrics and found similar results: a high correlation between MI and I_{sec} , and low correlation of I_{spike} with either MI or I_{sec} metrics (Fig. 7C left). Interestingly, correlation values increased after normalizing the metrics (Fig. 7C right), indicating that the standardized metrics capture similar spatial features.

Finally, we investigated the relation between information and firing rate separately for putative pyramidal cells and interneurons. As shown in Fig. 8A top panels, the range of information values for pyramidal cells and interneurons had greater overlap when estimated by MI, while the interneurons exhibited much lower information than pyramidal cells when estimated by I_{sec} and I_{spike} . Moreover, in the case of I_{spike} , the estimated information had a clear relation to firing rate. On the other hand, the normalized versions of each metric showed similar relations between firing rate and information (Fig. 8A bottom panels), resembling the relation between firing rate and percentage of correct decoding (Fig. 8B).

DISCUSSION

We studied three metrics of spatial information using both simulated and real data by comparing how well they correlate with the capacity of decoding the animal position from single-cell spike trains. Decoders are used to infer the most likely stimulus that elicited a particular response; their performance is directly linked to the information about the stimulus contained in the response (Quiroga and Panzeri, 2009). In other words, if two variables are related, it might be possible to use one of them to decode the other. The information contained in the confusion matrix of a decoder provides a lower bound to the information between the two variables (Quiroga and Panzeri, 2009), allowing the use of decoding performance as an empirical estimate of the real (spatial) information of the cell (Robertson et al., 1999; Jensen and Lisman, 2000; Huxter et al., 2008; Lopes-dos-Santos et al., 2015). In this work, the performance of a Bayesian decoder was assumed to represent the gold standard of the true spatial information content of a cell. Although more complex decoders could potentially present better performance, the naïve Bayes approach was chosen due to its efficiency and simplicity. Under this framework,

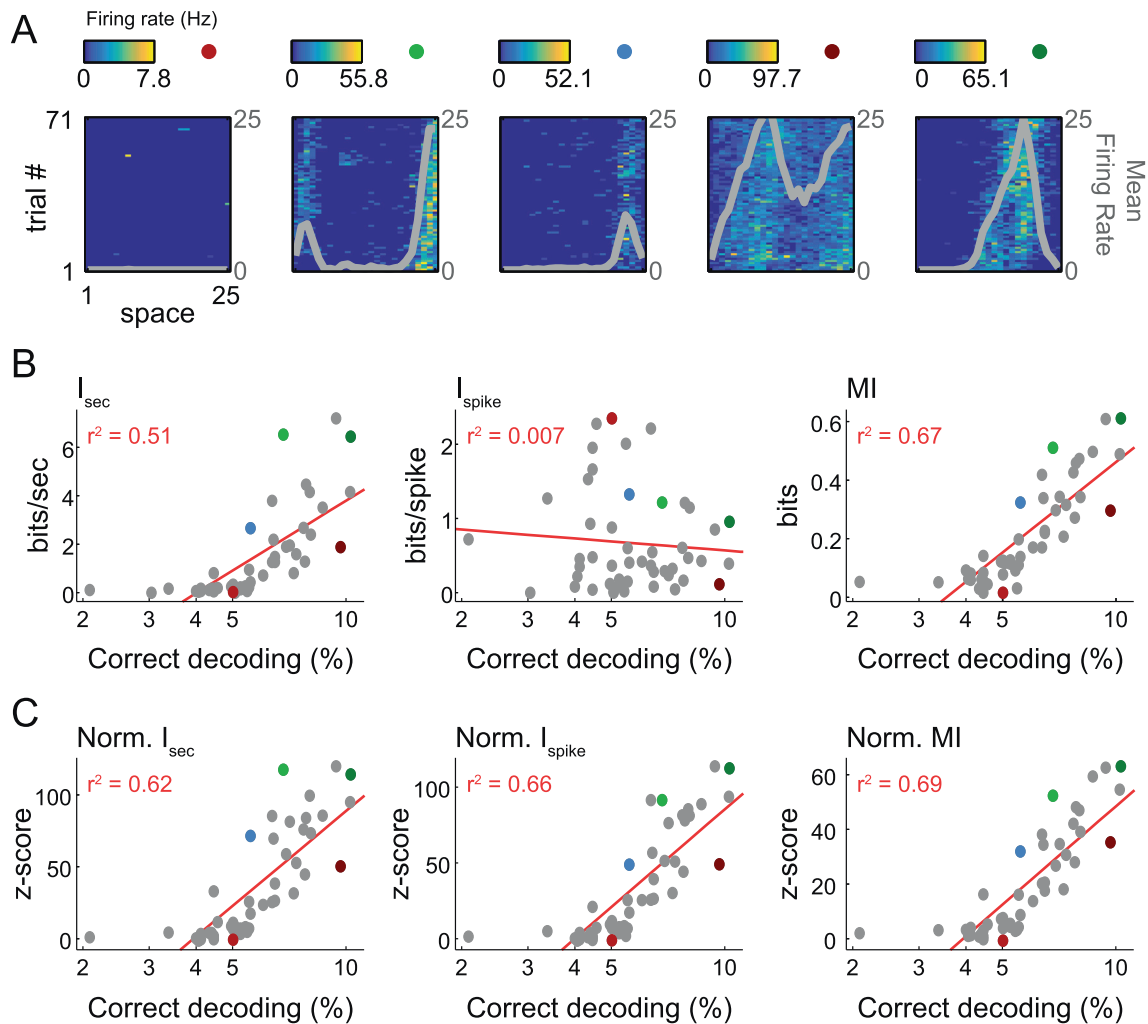


Fig. 5. Spatial information metrics applied to real cell data. (A) Example of the spatial activity of five neurons recorded in hippocampal CA1 of one rat during a linear track session. The recording included spatially modulated (i.e., place cells) as well as non-modulated cells. Colored circles mark the same cells highlighted in B and C. (B) Scatter plots of information estimates and decoding performance in log scale for all recorded cells. Note that the MI exhibits the best correlation with decoding performance. (C) Same as in B but for the normalized metrics. As in simulated data (Figs. 3 and 4), normalizing I_{sec} and I_{spike} improves their correlation with decoding performance.

our simulations show that the MI better correlates with spatial decoding performance than I_{sec} and I_{spike} . Similar results hold when running our analyses on real linear-track data (Figs. 5 and 6), though I_{sec} had better correlation with correct decoding in real than simulated data. This might be explained by the absence of some types of spatial modulation considered in our simulated cells, along with the expected presence of classical place cells in CA1 recordings. Finally, we found that the correlation with decoding performance achieves the same levels among the metrics following a surrogate-based normalization.

The low correlation of I_{sec} and I_{spike} with correct decoding in simulated data is a consequence of the way these metrics quantify information. Because they use the average firing rate over trials, a single trial with high firing rate can bias the metrics toward higher information values. This issue was apparent for simulated cells that were spatially modulated only in some of the trials (cell types IV and V): cells with different numbers of modulated trials but same mean firing rate showed

similar information values (trial consistency; see Fig. 9). This contrasts with the intuitive notion that the more consistent the spatial modulation of a cell across trials, the higher its spatial information. Another characteristic of I_{sec} and I_{spike} metrics was their sensitivity to changes in basal firing rate. For instance, the same increase in firing rate but from different baseline levels (e.g., 0 to 5 Hz vs. 10 to 15 Hz) leads to different information values, favoring cells with low basal firing rate to have higher information (additive effect; Fig. 9). Additionally, the I_{sec} metric is sensitive to changes in the mean firing rate of a cell upon a multiplicative factor (multiplicative effect; Fig. 9).

Some of these effects may be due to the assumptions underlying the derivation of I_{sec} and I_{spike} . Interpreting the information rate in bits per second may lead one to consider that this rate is valid for any amount of time. However, the estimated information rate might be valid only for small Δt , during which the cell emits only one or no spike. Since the information conveyed by a spike is

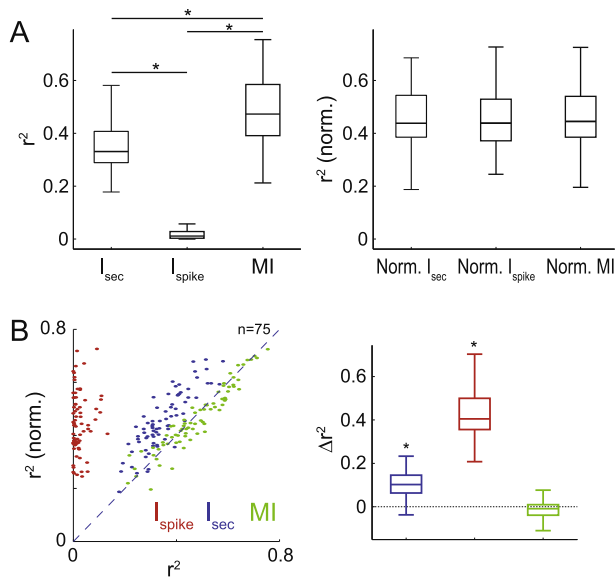


Fig. 6. Correlation of spatial metrics and decoding performance. (A) Boxplots show the distribution of coefficients of determination (r^2) between the logarithm of the percentage of correct decoding and I_{sec} , I_{spike} or MI for 75 linear track sessions. The original and normalized metrics are shown in the left and right panels, respectively. Notice similar r^2 values for the normalized metrics. $p < 0.001$ (Wilcoxon's signed-rank tests, Bonferroni's corrected). (B) Scatter plot of the r^2 for each metric and its normalized version (left) and boxplots of the changes in r^2 after normalization (right). There is a significant increase in r^2 for I_{sec} and I_{spike} ; $p < 0.001$ (Wilcoxon's signed-rank test, Bonferroni's corrected).

not independent of the information from previous ones, the information rate may be overestimated as Δt increases (Skaggs et al., 1993). The amount of redundancy is then directly related to the firing rate of the cell, explaining why normalizing I_{sec} by the firing rate (i.e., I_{spike}) impedes the overestimation caused by the multiplicative effect (Fig. 9). The underestimation of the additive effect, on the other hand, is due to the fact that the firing rate of each spatial bin becomes closer to the average firing rate; that is, adding a constant to all firing rate bins reduces the variations across bins, making the ratio λ_i/λ closer to 1 and $\log(\lambda_i/\lambda)$ closer to 0. This effect is inherent to the definition of both I_{sec} and I_{spike} , albeit the convergence to zero is faster for I_{spike} (Fig. 9).

In contrast to I_{sec} and I_{spike} , the MI was able to capture the true spatial information of the cells irrespective of rate-position map features. For instance, in the example cases shown in Fig. 3, the MI was capable of properly estimating the information despite differences in mean firing rate (cell types I and II). This is because the MI is based on the probability of each firing rate value and not on the value itself. In other words, as opposed to I_{sec} and I_{spike} , the MI is insensitive to additive and multiplicative effects (Fig. 9). Moreover, the MI detected the increase in spatial information with higher trial consistency despite the constant mean firing rate across trials (cell type V). Notice that the MI takes into account the firing rate of every trial instead of the trial mean, making this metric more robust to inter-trial variability.

Using information-theoretic metrics has been a frequent strategy to investigate neuronal encoding in different brain areas. For instance, the MI has been previously applied to a variety of experimental paradigms aimed at identifying the neuronal correlates of sensory and motor functions as well as of high-level processes such as working memory (Panzeri et al., 2001; Petersen et al., 2001; Belitski et al., 2008, 2010; Montemurro et al., 2008; Kayser et al., 2009; Pavão et al., 2014; Lopes-dos-Santos et al., 2015; Rossi-Pool et al., 2016; Vergara et al., 2016). Moreover, the MI has also been used to study diverse coding strategies, such as the amount of information carried by different types of signals (e.g., LFP oscillations and spikes) and/or by a combination of their features (Belitski et al., 2008, 2010; Montemurro et al., 2008; Kayser et al., 2009; Pavão et al., 2014). Estimating the amount of information a stochastic process can convey about itself (entropy) and another variable (MI) is a useful tool to study unknown relations between different signals, which are often non-linear. Under the MI framework, virtually any combination of variables (e.g., continuous and discrete) from various sources (e.g., behavioral and physiological) can be studied with the only need of estimating their probability density functions. In our case, the underlying relationship between the animal position (a behavioral variable) and the neuronal firing rate (a physiological variable) could be accessed by the MI, which yielded information estimates that correlated well with decoding accuracy of a Bayesian classifier. Along with previous literature (Quiroga and Panzeri, 2009), this result illustrates how versatile Information Theory is and the relevance of the standard MI definition for studying neuronal encoding.

We introduced a normalization that estimates and corrects for the intrinsic bias present in the rate-position map. This bias is due to the fact that even shuffled firing rate maps will have information estimates above zero (i.e., random firing rate maps are seldom constant over space). After normalizing I_{sec} and I_{spike} by the mean and standard deviation of the shuffled distribution, we found an increase in their correlation with decoding performance, which reached the same level as the original MI (Fig. 4). These normalizations may be useful under situations in which the definition of a trial is not possible or the probabilities underlying the MI cannot be easily computed. It is worth noticing that binning the firing rate values to compute the MI might introduce information bias (Panzeri et al., 2007), which can also be corrected by the normalization.

These findings may have implications on what we know about spatial representations in the brain. For instance, previous studies measured spatial information in entorhinal cortex and hippocampus and found higher spatial information in putative excitatory than inhibitory neurons (Frank et al., 2000, 2001). However, this could reflect the bias of I_{spike} toward higher information values in cells with low mean firing rate (Fig. 9B). Accordingly, we found that information estimates for putative pyramidal cells and interneurons in CA1 drastically differ for I_{spike} but not for MI (Fig. 8). Moreover, our results further show that the three metrics can yield different subpopulations of

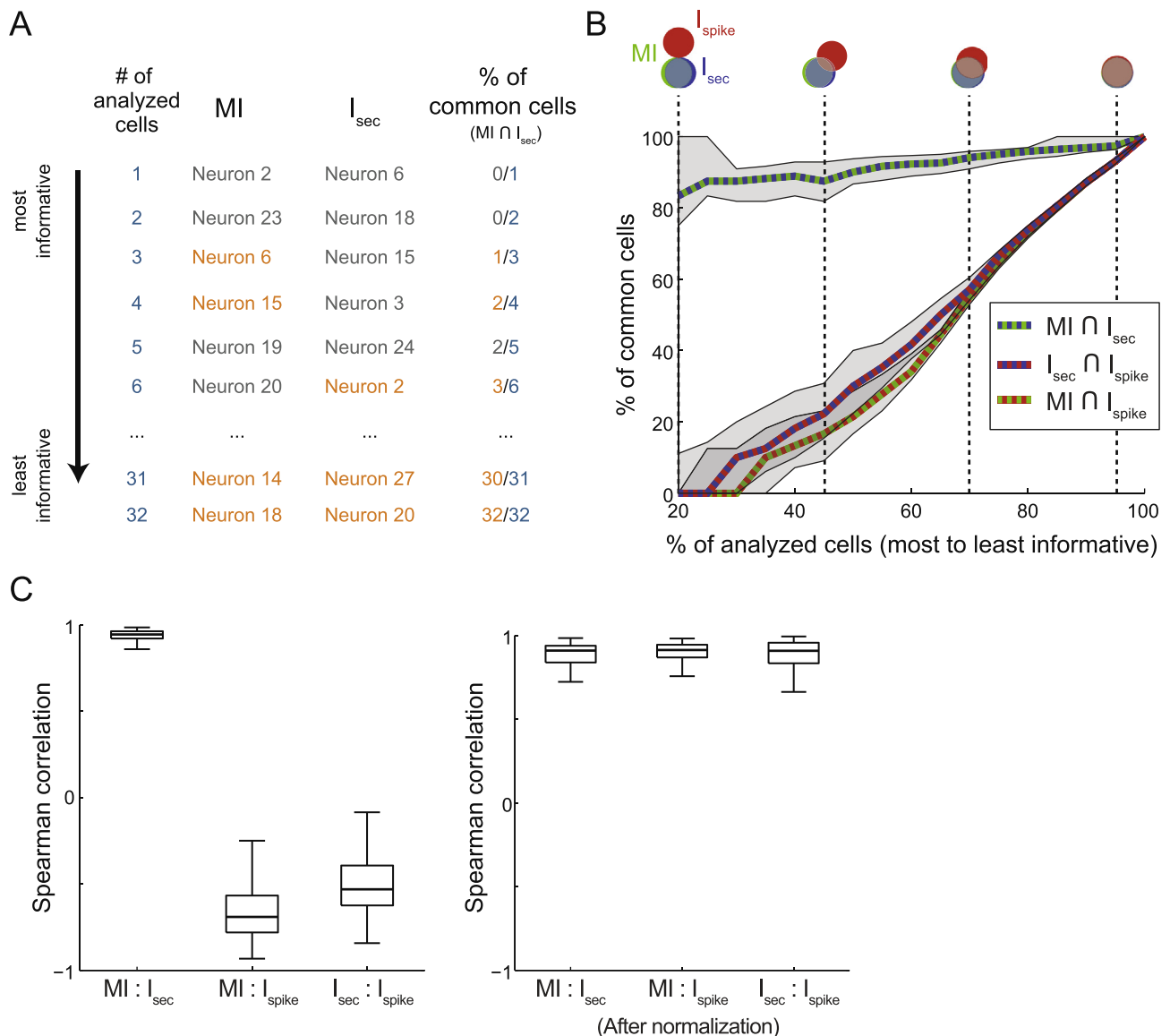


Fig. 7. Comparison of the subpopulations of spatial cells according to each metric. (A) For each linear track session, neurons were ranked from most to least informative according to I_{sec} , I_{spike} or MI. The intersections (% of common cells) between the top informative neurons of each metric pair was computed considering the first N cells, with N varying from 1 to the total number of cells. The panel illustrates this procedure for a pair of metrics (MI and I_{sec}) in one session. (B) Median % of common cells plotted against the number of analyzed cells (in percentage from total cell number, N%) across all sessions for each pair of metrics. The shaded area represents interquartile range. Venn diagrams on top show the median intersection at the corresponding N%. Note the low intersection between I_{sec} and I_{spike} or MI and I_{spike} . (C) Boxplot distributions of the Spearman correlation between each pair of original (left) and normalized (right) metrics.

spatially modulated neurons (Fig. 7). Surprisingly, despite their similar definition, we found very low intersection between neurons classified using I_{spike} and I_{sec} (Fig. 7B).

More generally, our observations raise the question of what defines the spatial information of a neuron. For instance, most of the known correlates of space in the hippocampus focus on neurons that typically spike at very low rates outside their receptive fields (O'Keefe and Dostrovsky, 1971; Taube et al., 1990; Fyhn et al., 2004; Hafting et al., 2005). While I_{spike} and (mainly) I_{sec} work well to select cells with high spatial information in these cases, we wonder whether these metrics may have limited our understanding of spatial coding. In other

words, the I_{sec} and I_{spike} are well suited to detect “canonical” place cells (i.e., cells whose spatial firing rate is a unimodal function centered on the place field), but may fail to detect other spatially informative neurons whose firing rate maps are not that of a canonical place cell. Notice that adding a constant factor to a spatial firing rate map does not influence decoding performance nor the information estimated by the original MI, but decreases the amount of spatial information estimated by I_{sec} and I_{spike} (Fig. 9). While some researchers may intuitively consider that cells which are silent outside the place field convey more information when they spike than cells with high basal firing rates, others may be more concerned as to

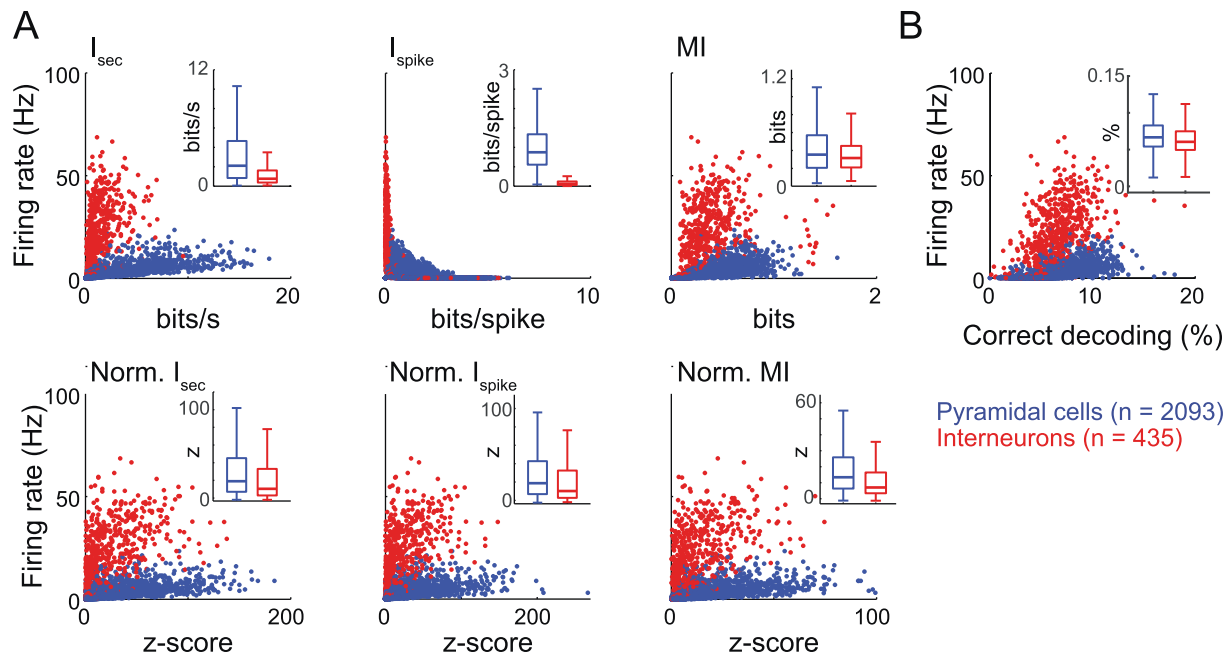


Fig. 8. Information estimates for CA1 pyramidal cells and interneurons. (A) Scatter plots of spatial information and firing rate for each metric (top) and its normalized version (bottom). Only neurons classified as pyramidal cells (blue; $n = 2093$) and interneurons (red; $n = 435$) were considered. Insets show boxplot distributions of information values. (B) Scatter plots of decoding performance and firing rate. Note inverse dependence of I_{spike} on firing rate, and greater overlap in the information range of interneurons and pyramidal cells for MI and the normalized metrics, akin to the overlap present in decoding performance. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

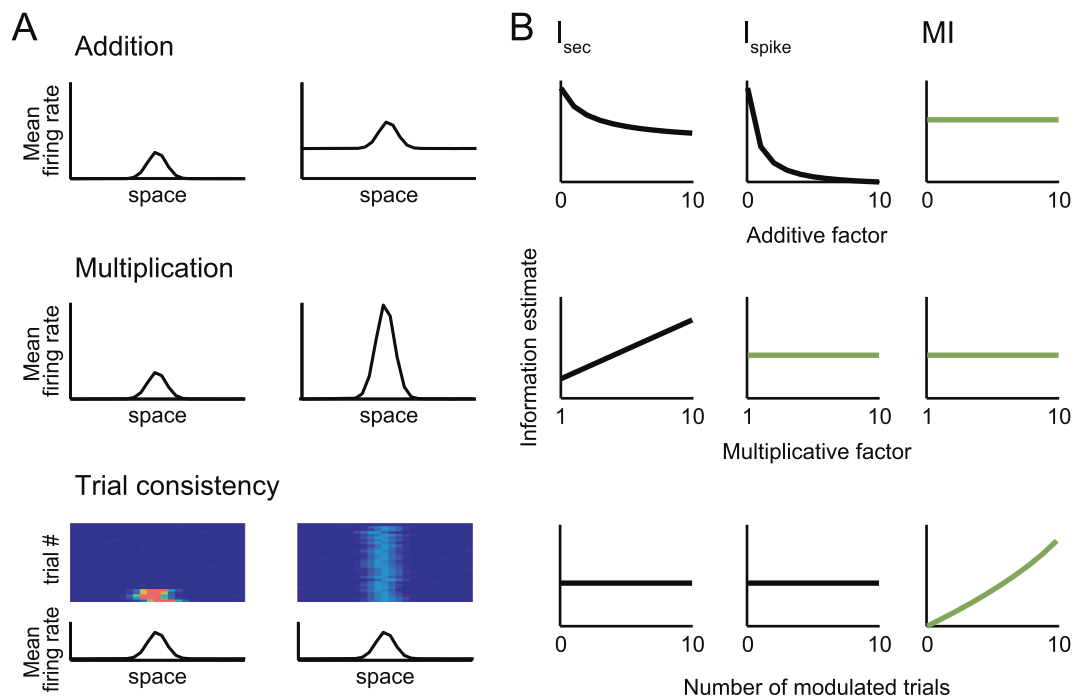


Fig. 9. Influence of addition, multiplication and trial consistency on spatial metrics. (A) Schematic examples of firing rate addition (top) and multiplication (middle), as well as of consistency over trials (bottom). In the latter case, the mean firing rate over trials was fixed. (B) I_{sec} , I_{spike} and MI for different parameters of the cases in A. I_{sec} and I_{spike} decay as the additive factor increases, while I_{sec} linearly increases with the multiplication factor. Only the MI increases with trial consistency. The green curves are cases in which the metric behaves similar to the decoding performance.

whether it is possible or not to decode the animal position from the firing rate of the cell. The latter would further argue that every cell that carries information about the

animal position (as retrieved by a decoder) could be called a “place cell”, independently of place field shape or mean firing rate. Noteworthy, the existence of uncon-

ventional spatial representations has been recently demonstrated for medial entorhinal cortex neurons (Diehl et al., 2017; Hardcastle et al., 2017a; see also Hardcastle et al., 2017b). In any case, our results suggest that the MI is a suitable metric to capture other types of less canonical spatial correlates that may have gone undetected so far. In cases where the MI cannot be computed, the normalization of I_{sec} and I_{spike} may constitute good alternatives.

ACKNOWLEDGMENTS

This work was supported by CAPES and CNPq, Brazil. We thank the Buzsáki laboratory for making data publicly available at <http://crcns.org/>, a data-sharing website supported by NSF and NIH, USA. We thank D. Laplagne and R. Romcy-Pereira for critical discussions.

REFERENCES

- Barbieri R, Wilson MA, Frank LM, Brown EN (2005) An analysis of hippocampal spatio-temporal representations using a Bayesian algorithm for neural spike train decoding. *IEEE Trans Neural Syst Rehabil Eng* 13:131–136.
- Belitski A, Gretton A, Magri C, Murayama Y, Montemurro MA, Logothetis NK, Panzeri S (2008) Low-frequency local field potentials and spikes in primary visual cortex convey independent visual information. *J Neurosci* 28:5696–5709.
- Belitski A, Panzeri S, Magri C, Logothetis NK, Kayser C (2010) Sensory information in local field potentials and spikes from visual and auditory cortices: time scales and frequency bands. *J Comput Neurosci* 29:533–545.
- Brown EN, Frank LM, Tang D, Quirk MC, Wilson MA (1998) A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. *J Neurosci* 18:7411–7425.
- Diehl GW, Hon OJ, Leutgeb S, Leutgeb JK (2017) Grid and non-grid cells in medial entorhinal cortex represent spatial location and environmental features with complementary coding schemes. *Neuron* 94:83–92.
- Ego-Stengel V, Wilson MA (2007) Spatial selectivity and theta phase precession in CA1 interneurons. *Hippocampus* 17:161–174.
- Eichenbaum H (2000) A cortical–hippocampal system for declarative memory. *Nat Rev Neurosci* 1:41–50.
- Frank LM, Brown EN, Wilson M (2000) Trajectory encoding in the hippocampus and entorhinal cortex. *Neuron* 27:169–178.
- Frank LM, Brown EN, Wilson MA (2001) A comparison of the firing properties of putative excitatory and inhibitory neurons from CA1 and the entorhinal cortex. *J Neurophysiol* 86:2029–2040.
- Fyhn M, Molden S, Witter MP, Moser EI, Moser M-B (2004) Spatial representation in the entorhinal cortex. *Science* 305:1258–1264.
- Hafting T, Fyhn M, Molden S, Moser M-B, Moser EI (2005) Microstructure of a spatial map in the entorhinal cortex. *Nature* 436:801–806.
- Hardcastle K, Maheswaranathan N, Ganguli S, Giocomo LM (2017a) A Multiplexed, heterogeneous, and adaptive code for navigation in medial entorhinal cortex. *Neuron* 94:375–387.
- Hardcastle K, Ganguli S, Giocomo LM (2017b) Cell types for our sense of location: where we are and where we are going. *Nat Neurosci* 20:1474–1482.
- Huxter JR, Senior TJ, Allen K, Csicsvari J (2008) Theta phase-specific codes for two-dimensional position, trajectory and heading in the hippocampus. *Nat Neurosci* 11:587–594.
- Jensen O, Lisman JE (2000) Position reconstruction from an ensemble of hippocampal place cells: contribution of theta phase coding. *J Neurophysiol* 83:2602–2609.
- John GH, Langley P (1995) Estimating continuous distributions in Bayesian classifiers. In: *Proceedings of the eleventh conference on uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., p. 338–345.
- Kayser C, Montemurro MA, Logothetis NK, Panzeri S (2009) Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. *Neuron* 61:597–608.
- Kropff E, Carmichael JE, Moser M-B, Moser EI (2015) Speed cells in the medial entorhinal cortex. *Nature* 523:419–424.
- Lopes-dos-Santos V, Panzeri S, Kayser C, Diamond ME, Quiroga RQ (2015) Extracting information in spike time patterns with wavelets and information theory. *J Neurophysiol* 113:1015–1033.
- Mizuseki K, Sirota A, Pastalkova E, Buzsáki G (2009) Theta oscillations provide temporal windows for local circuit computation in the entorhinal–hippocampal loop. *Neuron* 64:267–280.
- Mizuseki K, Sirota A, Pastalkova E, Diba K, Buzsáki G (2013), Multiple single unit recordings from different rat hippocampal and entorhinal regions while the animals were performing multiple behavioral tasks. *CRCNS Org*.
- Montemurro MA, Rasch MJ, Murayama Y, Logothetis NK, Panzeri S (2008) Phase-of-firing coding of natural visual stimuli in primary visual cortex. *Curr Biol* 18:375–380.
- Morris RGM, Garrud P, Rawlins JNP, O'Keefe J (1982) Place navigation impaired in rats with hippocampal lesions. *Nature* 297:681–683.
- Moser EI, Kropff E, Moser M-B (2008) Place cells, grid cells, and the brain's spatial representation system. *Annu Rev Neurosci* 31:69–89.
- O'Keefe J, Dostrovsky J (1971) The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res* 34:171–175.
- O'Keefe J, Recce ML (1993) Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus* 3:317–330.
- Panzeri S, Petersen RS, Schultz SR, Lebedev M, Diamond ME (2001) The role of spike timing in the coding of stimulus location in rat somatosensory cortex. *Neuron* 29:769–777.
- Panzeri S, Senatore R, Montemurro MA, Petersen RS (2007) Correcting for the sampling bias problem in spike train information measures. *J Neurophysiol* 98:1064–1072.
- Pavão R, Piette CE, Lopes-dos-Santos V, Katz DB, Tort AB (2014) Local field potentials in the gustatory cortex carry taste information. *J Neurosci* 34:8778–8787.
- Petersen RS, Panzeri S, Diamond ME (2001) Population coding of stimulus location in rat somatosensory cortex. *Neuron* 32:503–514.
- Quiroga RQ, Panzeri S (2009) Extracting information from neuronal populations: information theory and decoding approaches. *Nat Rev Neurosci* 10:173–185.
- Robertson RG, Rolls ET, Georges-François P, Panzeri S (1999) Head direction cells in the primate pre-subiculum. *Hippocampus* 9:206–219.
- Rossi-Pool R, Salinas E, Zainos A, Alvarez M, Vergara J, Parga N, Romo R (2016) Emergence of an abstract categorical code enabling the discrimination of temporally structured tactile stimuli. *Proc Natl Acad Sci U S A* 113:7966–7975.
- Salinas E, Romo R (1998) Conversion of sensory signals into motor commands in primary motor cortex. *J Neurosci* 18:499–511.
- Scoville WB, Milner B (1957) Loss of recent memory after bilateral hippocampal lesions. *J Neurol Neurosurg Psychiatry* 20:11–21.
- Shannon CE (1948) A mathematical theory of communication, Part I, Part II. *Bell Syst Tech J* 27:623–656.
- Skaggs WE, McNaughton BL, Gothard KM, Markus EJ (1993), An information-theoretic approach to deciphering the hippocampal code, in: *Advances in Neural Information Processing Systems* 5. Hanson, SJ, Giles, CL Cowan, JD (Eds.), pp. 1030–1037.
- Skaggs WE, McNaughton BL, Wilson MA, Barnes CA (1996) Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. *Hippocampus* 6:149–172.
- Taube JS, Muller RU, Ranck JB (1990) Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. *J Neurosci* 10:420–435.

- Vergara J, Rivera N, Rossi-Pool R, Romo R (2016) A neural parametric code for storing information of more than one sensory modality in working memory. *Neuron* 89:54–62.
- Zhang K, Ginzburg I, McNaughton BL, Sejnowski TJ (1998) Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells. *J Neurophysiol* 79:1017–1044.
- Zola-Morgan SM, Squire LR (1990) The primate hippocampal formation: evidence for a time-limited role in memory storage. *Science* 250:288–290.

(Received 15 September 2017, Accepted 31 January 2018)
(Available online 10 February 2018)